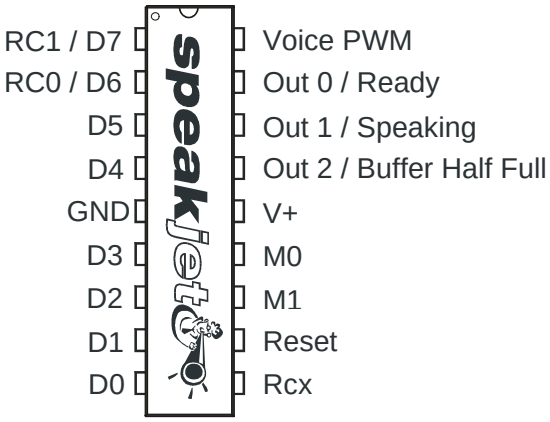
MAGNEVATION

speakjet

Natural Speech & Complex Sound Synthesizer

User's Manual

February 26, 2004



General Description

The SpeakJet is a completely self contained, single chip voice and complex sound synthesizer. It uses Mathematical Sound Architecture [tm] (MSA) technology to control an internal five channel sound synthesizer to generate on-the-fly, unlimited vocabulary speech synthesis and complex sounds. The SpeakJet is preconfigured with 72 speech elements (allophones), 43 sound effects, and 12 DTMF Touch Tones. Through the selection of these MSA components and in combination with the control of the pitch, rate, bend, and volume parameters, the user has the ability to produce unlimited phrases and sound effects, with thousands of variations, at any time.

The SpeakJet can be controlled simultaneously by logic changes on any one of its eight Event Input lines, and/or by a single I/O line from a CPU (such as the OOPic, Basic Stamp or PC) allowing for both CPU-Controlled and Stand-Alone operations.

Other features include an internal 64 byte input buffer, Internal Programmable EEPROM, three programmable outputs, and direct user access to the internal five channel sound synthesizer.

More information can be found at www.SpeakJet.com

Core Features:

- Programmable, 5 channel synthesizer.
- Natural phonetic speech synthesis.
- DTMF and other sound effects.
- Programmable control of pitch, rate, bend and volume.
- Programmable power – up or reset announcements.
- Multiple modes of operation.
- Simple interface to microcontrollers.
- Simple "Stand Alone" operation.
- Three programmable digital outputs.
- Internal 64 Byte buffer.
- Internal programmable EEPROM.
- Extremely low power consumption.
- Low pin count.
- Multiple case styles available.

Special Features

- Three multipurpose, programmable digital outputs allow the SpeakJet to control external devices based on timing of the sound output. Control of devices may include lights, motors, or even launch model rocket after a count down sequence.
- An internal clock oscillator provides for a truly "Self Contained sound system. Simply connect the SpeakJet to a power supply and a speaker to hear it speak. *
- An internal user programmable EEPROM allows for programming of up to 16 complex phrases or sound sequences. These may be played back once or looped many times in response to events.
- Phrases may call other phrases, sounds or controls, with nesting up to 3 levels deep.
- No special equipment is required to program the internal EEPROM, only a serial connection is required.

Interface Options

- **CPU Control:** Single Wire Serial Input from microprocessors such as the OOPic, Basic Stamp, or any other computer system equipped with a serial port.
- **Stand Alone:** Eight Event Inputs for execution of up to sixteen phrases, sound effects or control functions without a microcontroller.
- **RC Input:** Two Servo Inputs for execution of up to four phrases, sound effects or control functions via wireless model Airplane or Car Radio Control receivers.

SpeakJet Pin Details

Pin #	Description	Functional Details
1	RC1/D7	RC1 Input/ Event Input 7
2	RC0/D6	RC0 Input/ Event Input 6
3	D5	Event Input 5
4	D4	Event Input 4
5	GND	Ground
6	D3	Event Input 3
7	D2	Event Input 2
8	D1	Event Input 1
9	D0	Event Input 0
10	RCX	TTL Serial Input Active High
11	RESET	Master Reset Active Low
12	M1	Baud Select
13	M0	Demo Mode Active High
14	V+	Power input +2.0 to +5.5 volts DC
15	OUT2/Buffer Half Full	Output 2 (External)/Buffer Half Full (Internal)
16	OUT1/Speaking	Output 1 (External)/Currently Speaking (Internal)
17	OUT0/Ready	Output 0 (External)/SpeakJet Ready (Internal)
18	Voice PWM	8 KHz PWM Voice Output

Electrical Specifications:

Supply voltage	2.0 to 5.5 VDC
Supply Current:	
Idle:	<5ma. Plus loads
Speaking:	<5ma. Plus loads
Sink/Source Current:	
Outputs	25ma.
Inputs levels:	
High	Supply
Low	GND
EEPROM:	
Max. write cycles	Typical 1,000,000 times

Mechanical Specification:

Thermal storage:	-60 to +140 Degrees C
Thermal operating	-50 to +120 Degrees C

The thermal specifications are preliminary and may change as testing is completed.

SpeakJet Pin Functional Explanation

Event Input 0 – 7:

Event Inputs provides a way to execute phrases, sound effects or control functions that have been pre-programmed into the EEPROM. Each pin can be programmed to execute a phrase, sound effect or control function on a Low to High going transition, or a High to Low going transition. This can be combined to allow two phrases or sound effects to be played with one pin or disabled. This feature is controlled by an entry into the EEPROM. These are logical inputs that require a High or Low voltage level. When not used tie to GND

RC0, RC1 Input:

RC input allows the SpeakJet to connect to model radio control receivers and mimic a servo to execute up to four phrases, sound effects or control functions. These phrases, sound effects or control functions can be pre-programmed into the EEPROM. These are logical inputs that require a High or Low voltage level.

RCX:

Receive Serial Data Input. Gives external devices a means to communicate with the internal MSA ™, EEPROM and Control System by TTL serial data stream. This is a logical input that requires a High or Low voltage level.

RESET:

A low on this pin forces the internal control circuit to reset and clear the buffer. EEPROM stored phrases, sound effects, control functions and configuration bits are not affected. The use of RESET along with the logic states of the Mode Select inputs M0 and M1, places the SpeakJet into the different modes of operation. This is a logical input that requires a High or Low voltage level.

M0, M1:

Mode select pins places the SpeakJet into one of three modes: Normal Operation, Demo / Test mode, and Baud Rate Configure mode. These pins are by the internal control system after a power-up or a reset is forced. These pins are logical inputs that require a High or Low voltage level.

OUT2/Buffer Half Full:

This output line can be used for flow control with a serial port by connecting to the DSR line back to the computer or controller. It may also be used as a general purpose output line for controlling external devices such as indicator lights, piezo audio transducers, relays, MOSFET transistors, etc. This line is a logical output with either a High or Low voltage level. The function of this output line is stored in the EEPROM and the factory default configurations is: “Buffer Half Full” and is Active High.

OUT1/Speaking:

This output line can be used to indicate when the SpeakJet is "Speaking". This can be useful to switch on an amplifier or indicate that an audio output is present. It may also be used as a general purpose output line for controlling external devices such as indicator lights, piezo audio transducers, relays, MOSFET transistors, etc. This line is a logical output with either a High or Low voltage level. The function of this output line is stored in the EEPROM and the factory default configurations is: "Speaking" and is Active High

OUT0/Ready:

This output line can be used to indicate when the SpeakJet is Ready. This indicates that the internal self-test has passed and the Speakjet is "Ready" to accept data through one of its many inputs. It may also be used as a general purpose output line for controlling external devices such as indicator lights, piezo audio transducers, relays, MOSFET transistors, etc. This line is a logical output with either a High or Low voltage level. The function of this output line is stored in the EEPROM and the factory default configurations is: "Ready" and is Active High.

Voice PWM:

Voice PWM line is an output line with a square wave of 8KHz. The duty cycle of this line is varied by the modulation of the sound output. This duty cycles varies by as much as approximately 20% to 80% depending on the level of output represented by the audio wave. A simple "two pole" low pass filter is all that is required before an amplifier to obtain high quality sound output. See Figure 1

Mode Select:

When a High logic level is placed on Mode Select, M0, and RESET is forced, the SpeakJet starts a Demo mode where all internal allophones and sound effects are played in a loop. This loop plays if the buffer is empty and will suspend until the buffer is empty again if data is sent to the buffer. When M0 is brought back to a low state the demo mode is suspended. If M0 is brought back to a high state, the demo mode continues from the point where it was suspended. A RESET with a low logic level on Mode select M0 will discontinue the Demo/Test mode and place the SpeakJet back to normal operation.

If the SpeakJet is in the Demo/Test mode and a momentary Low logic level is placed on Mode Select M1, the Speakjet starts a Sonar Ping sound to indicate the Baud Rate Configuration mode is selected. At this time the Speakjet is waiting for a serial sync character (hex 55) to set the internal clock for the baud rate that is to be used. Once the baud rate is captured, the Speakjet automatically stores value this value in the EEPROM for future reference even after a power down condition and returns to the demo mode.

Normal operation requires that M0 to be Low logic level and M1 to be a High logic level.

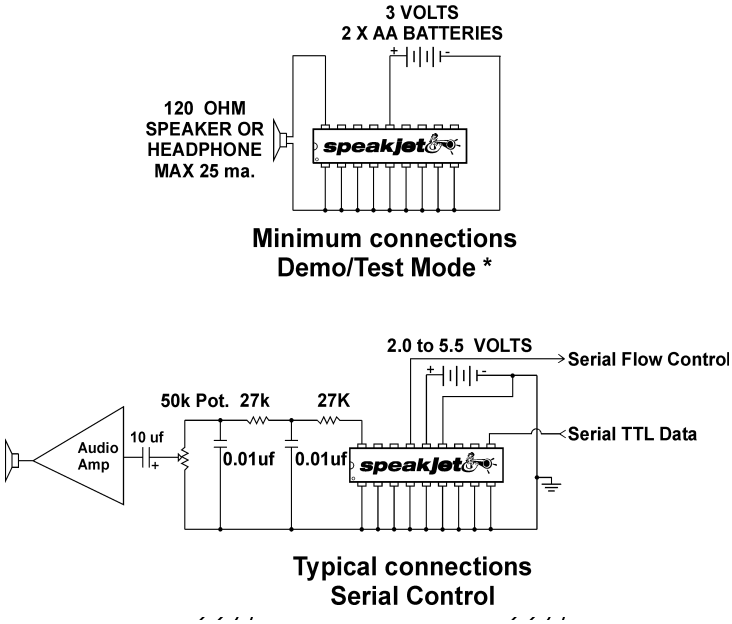


Figure 1. SpeakJet Typical connections

Serial Data:

Serial Data is the main method of communicating with the SpeakJet. By utilizing an internal Baud rate generator with an Auto Baud Rate Detector, a wide range of serial data can be read and interpreted to execute voices, sounds or commands. The serial data can also be used to program the internal EEPROM. The SpeakJet does not accept actual RS-232A signal levels and will be damaged if attempted to input these levels. In order to read data from a RS-232A type of serial data stream, a level shifter/line receiver must be used. This can be as simple as a small transistor or more robust device like the MAX232A from Maxim.

The SpeakJet serial configuration is fixed at : Non Inverted, No-Parity, 8 bits and 1 stop bit (N,8,1). The Speakjet can accept Baud rates from 300 to 19200. The factory default setting is 9600 baud. This baud rate can be changed any time by placing the SpeakJet into the "Baud Rate Configure Mode". The Speakjet may need to have the "Baud Rate Configure Mode" executed if it is placed in an area where the temperature has changed greatly from where it was last performed. This is due to the internal Oscillator drift with a large temperature change.

The Speakjet utilizes SCP (Serial Control Protocol) developed by and licensed from Savage Innovations – Makers of the OOPic, to give the end user access to all of the low level registers for the Control, EEPROM, Synthesizers, and the digital Outputs.

Preliminary Data

General Operation:

The SpeakJet is an advanced CMOS LSI device that incorporates six basic internal modules: (Figure 2)

- 5-Channel Synthesizer. For generating both voice and sound effects.
- Sound Component Database and Mathematical Sound Architecture ^{tm} (MSA) engine that controls the 5-Channel Synthesizer.
- Chip control I/O, which consists of 1 Serial input line, 8 Event input lines, 3 output lines, and 2 mode select input lines.
- 64-Byte input buffer.
- User EEPROM.
- A Serial Control Protocol (SCP) supervises the incoming serial commands.

When powered on, the SpeakJet will first read the Mode Select line M0 to determine if the SpeakJet is to operate in Demonstration Mode. Once the Initial Mode is set, (note that this only occurs at power-up) the input buffer is cleared, the status lines are set and the SpeakJet begins operating and the ready line is brought high.

During operation, the SpeakJet waits for incoming command data, which specifies which sounds the SpeakJet is to produce and how to produce them.

The commands sent to the SpeakJet can be any of 256 commands. There are 7 operational groups of commands. SCP, Allophones, Sound Effects, DTMF, Pauses, Levels and Controls. Each operational group will behave in different ways.

With the exception for SCP commands, each command received, is buffered into a 64-Byte input buffer and executed in the order that they are received, First In First Out (FIFO). In the case of the SCP, these commands are executed immediately as they are received and not stored in the input buffer. See the section on SCP for further details.

As the SpeakJet executes the commands stored in the input buffer, each command is removed from the input buffer after the command is executed.

Both the Serial Input line and the 8 Event input lines provide methods of placing command data in the input buffer.

Serial Input Line: a single input line is used to receive command data at the preset baud rate. The Rcx line is configured as the serial input line and is expecting a TTL level (non inverted) signal.

Event Input Lines: In the case that the event line is configured to do so, a single EEPROM Call command is stored into the input buffer whenever one of the Event lines logic state is changed or in the case of the RC/TTL lines, when a PWM pulse changes a predetermined amount.

Status lines provide information on the current operational status of the SpeakJet.

“Ready” is active high when the SpeakJet is on and ready to accept commands.

“Speaking” is active high when the SpeakJet is currently generating any sound.

“Buffer Half Full” is active high when the input buffer is 1/2 full. A Buffer Half Full status indicates when the 64-Byte input buffer has more than 32 bytes in it. As the SpeakJet executes the commands stored in the input buffer, the buffer is freed and as 32 bytes of the buffer becomes available. The Buffer Half Full status line will reflect that status. This line can be used in a serial connection as the serial port’s RTS line. A logical Low output indicates that the buffer can accept 32 bytes and a logical High output indicates that the buffer cannot.

The status lines can also be user programmed to control external devices. Each Status line can be individually selected to be and internal Status or user selected logic state.

SCP:

Preliminary Data

ALLOPHONE SPEECH SYNTHESIS PRIMER

Introduction:

The sounds that are used to form the words of any particular language are called phonemes. During the articulation of the words of that language, variations of phonemes called allophones are vocalized one after another producing the spoken words and Phrases. The allophone speech synthesis technique provides the ability to synthesize a reproduction of that language by providing a method of playing allophone like sounds one after another in much the same way that the human mouth does.

The SpeakJet accomplishes allophone speech synthesis via Mathematical Sound Architecture [tm] (MSA). The function of MSA is to mathematically model the sounds that the human mouth produces as it moves from one position to the next. Seventy-Two discrete speech sound elements (one for each allophone) are stored in the SpeakJet's internal MSA Sound Component Database. Each MSA allophone component was created to duplicate a specific allophone sound made by the human mouth during English speech.

To produce speech, a list of selected allophones is sent to the SpeakJet. As the SpeakJet is vocalizing this list of allophones, MSA actively and continuously calculates all the sound components of the allophones including the transitional sounds made between the allophones producing the same in-between allophone sounds that the human mouth does.

Any English word or phrase can thusly be created by selecting the appropriate combination of allophones and pauses. Further tuning with the Rate, Pitch, Bend and Volume parameters adds to the delivery of the phrase and can change the emotion in which the phrase is perceived.

Stressing the Rate, Pitch, Bend and Volume parameters to levels outside the human range can result in some interesting sounds that go way beyond what a normal human mouth can produce. In addition, several other sounds, which are included in the MSA Sound Component Database, of which, some use vocalization modeling and some do not, can be integrated into the phrases.

The result is a system that gives the user the ability to not only produce an unlimited vocabulary, but also to produce slang, gibberish, moans, groans, yodels and other weird vocalized sounds not normally included in a canned TTS system.

Producing English Speech:

Five basic linguistic concepts will help you to create highly intelligible sounding speech with the SpeakJet.

- First, there is no one-to-one correspondence between written letters and speech sounds. Each sound in a language may be represented by more than one letter and, conversely each letter may represent more than one sound. (See the examples in Table A.) Because of these spelling irregularities, it is necessary to think in terms of sounds, not letters, when creating phrases.

- Secondly, speech sounds are acoustically different depending upon their position within a word. For example, the initial B sound in "Beep" is acoustically different from the B sound in "Box" The B sounds differ due to the influence of the vowels which follow them.

- Third, the human ear may perceive the same acoustic signal differently in the context of different sounds. Therefore, an allophone may sounds slightly different when used in various phrases.

- Fourth, some sounds in words are not actually pronounced and/or others are added when followed or preceded by certain other words. For example the two words “Night” and “Time”. Played separately, they both will articulate the “T” sound. However, played together they need to be articulated as “Ni-Time”.

- Fifth, The Vocalization Pitch, Play Rate and Frequency Bend parameters are just as important as the selection of the phonemes used. For instance, playing a phase that does not change the pitch at any point in time sounds very monotonic and robotic. While increasing the Pitch at the end of a sentence produces a questioning tone.

Phonemes of English Language:

Table A contains a chart of all the consonant phonemes of the English language. Table A also contains all the vowel phonemes.

Consonants are produced by creating an occlusion or constriction in the vocal tract, which produces an aperiodic sound source. If the vocal cords are vibrating at the same time, as in the case of the voiced fricatives VV, DH, ZZ, and ZH, (See Table 5) there are two sound sources: one which is aperiodic and one which is periodic.

Vowels are usually produced with a relatively open vocal tract and a periodic sound source provided by the vibrating vocal cords. They are classified according to whether the front or back of the tongue is high or low (See Table 4), whether they are long or short, and whether the lips are rounded or un-rounded. In English all rounded vowels are produced in or near the back of the mouth (UW, UH, OW, OH, OHRR, AW).

Preliminary Data

How to Use the Allophone Set:

Recall that a phoneme is acoustically different depending upon its position within a word. Each of these positional variants is an allophone of the same phoneme. An allophone, therefore, is the manifestation of a phoneme in true speech signal.

The following is a brief tutorial on creating speech using the SpeakJet allophone set.
Note that these are suggestions, not rules.

The allophone set (Refer to Table A) references the variations of each phonemes of the English language and give suggestions for the appropriate times that they are used. For example, there are 4 B sounds. “BE” and “BO” sound good in initial position, as in the words "Beep" and "Box" while “EB” and “OB” sound good in final position, as in "Rib" and "Club". A vowel modifier is also associated with the “BE”, “BO”, “EB” and “OB” sounds. The BE is used when followed by a front vowel sound and the BO is used when followed by a back vowel sound. Note that either can be used when followed by a central vowel sound. Also note that a B sound in the middle of a word can be either a initial position or a final position sound. For example, the word rabbit; is it RAB-IT or RA-BIT? Local dialect can be the deciding factor in cases like this.

Typically the first allophone in a phrase is longer than the rest of the allophones used with the phrase. Therefore, to create an initial SS, you can use the SLOW function before the SS allophone which will cause the SS sound to play for 1 and 1/2 the normal speed. If an even longer sound is needed then the SS allophone can be played twice. (Note that this cannot be done with all allophones.)
Repeating Diphthongs for example
will cause the diphthong slide to play twice which will not result in just a longer sound.

Diphthongs are perceived as single vowel sounds, but in reality they are sliding sounds that transition from one vowel sound into another vowel sound. For example the "A" sound in "Make" isn't a single vowel sound. It is a diphthong that transitions from the "EY" sound to the "IY" sound and thusly the name for the diphthong is "EYIY". This sliding sound is repeated when diphthongs are repeated. For example: "EYIY, EYIY" does not have the same affect as "IY, IY". Two "IY"s played back to back will produce an "IY" sound that lasts twice as long while the "EYIY, EYIY" will just repeat the slide. Five of the "R" sounds are diphthongs that begin with a vowel and end with the "R" sound. For example, the "AWRR" in "alarm" and the "OHRR" in "score". Determining when to use a diphthong and when to use the independent sound again largely relies on person preference and local dialect.

Some sounds (P, B, T, D, K, G, CH, and JH) require a brief duration of silence before them. For these, the silence has already been added but you may decide you want to add more. To increase the duration of silence the "SLOW" command can be used and optionally, a pause can be inserted before it, There are several different types and lengths of pauses that can be inserted in front of an allophone but be aware that different pauses will affect the way the MSA calculates the transitions from the last allophone, thought the silence and into the next allophone which may not produce the desired flow of sounds.

Stress can be accomplished in two ways. One is to cause vowels to play for a longer period of time. For example, in the word "extent" use the "Fast" command in front of the "EH" in the first syllable, which is unstressed and a "SLOW" command, or and additional "EH" in front of the "EH" in the second syllable which is stressed. A second way is to preceded the allophone with the "STRESS" and RELAX commands. The STRESS command duplicates the affect of slightly contracting the muscles of the mouth and the relax command duplicates the affects of slightly relaxing the muscles of the mouth. For example; "STRESS, IH" sounds more like (but not quite) the "IY" sound. Likewise, "RELAX, IY" sounds more like (but not quite) an "IH" sound. Note that if you elect to use the "STRESS" or "RELAX" command in combination with two "EH"s, then two command will be needed, one in front of each of the "EH"s.

Remember that you must always think about how a word sounds, not how it is spelled. For example, The N sound in "Link" is actually the "NGE" sound and the sounds in the words "letter" and “little" use the allophones "AXRR" and "EHEL". some sounds may not even be represented in words by any letters, as the "IYUW" in "computer”.

You will want to experiment with all the possible sounds to discover which version works best in any particular cluster of allophones.

One last suggestion is that you may want to add a short pause between words and a long pause between clauses.

Preliminary Data

Terms

pho·neme n.
The smallest phonetic unit in a language that is capable of conveying a distinction in meaning, as the M of Mat and the B of Bat in the English language.

al·lo·phone n.
A predictable phonetic variant of a phoneme. For example, the aspirated T of Top, the unaspirated T of sTop, and the TT (pronounced as a flap) of baTTer are allophones of the English phoneme /T/.

syl·la·ble n.
A unit of spoken language consisting of a single uninterrupted sound formed by a vowel, diphthong, or syllabic consonant alone, or by any of these sounds preceded, followed, or surrounded by one or more consonants.

syn·the·sis n.
The combining of separate elements or substances to form a coherent whole.

Preliminary Data

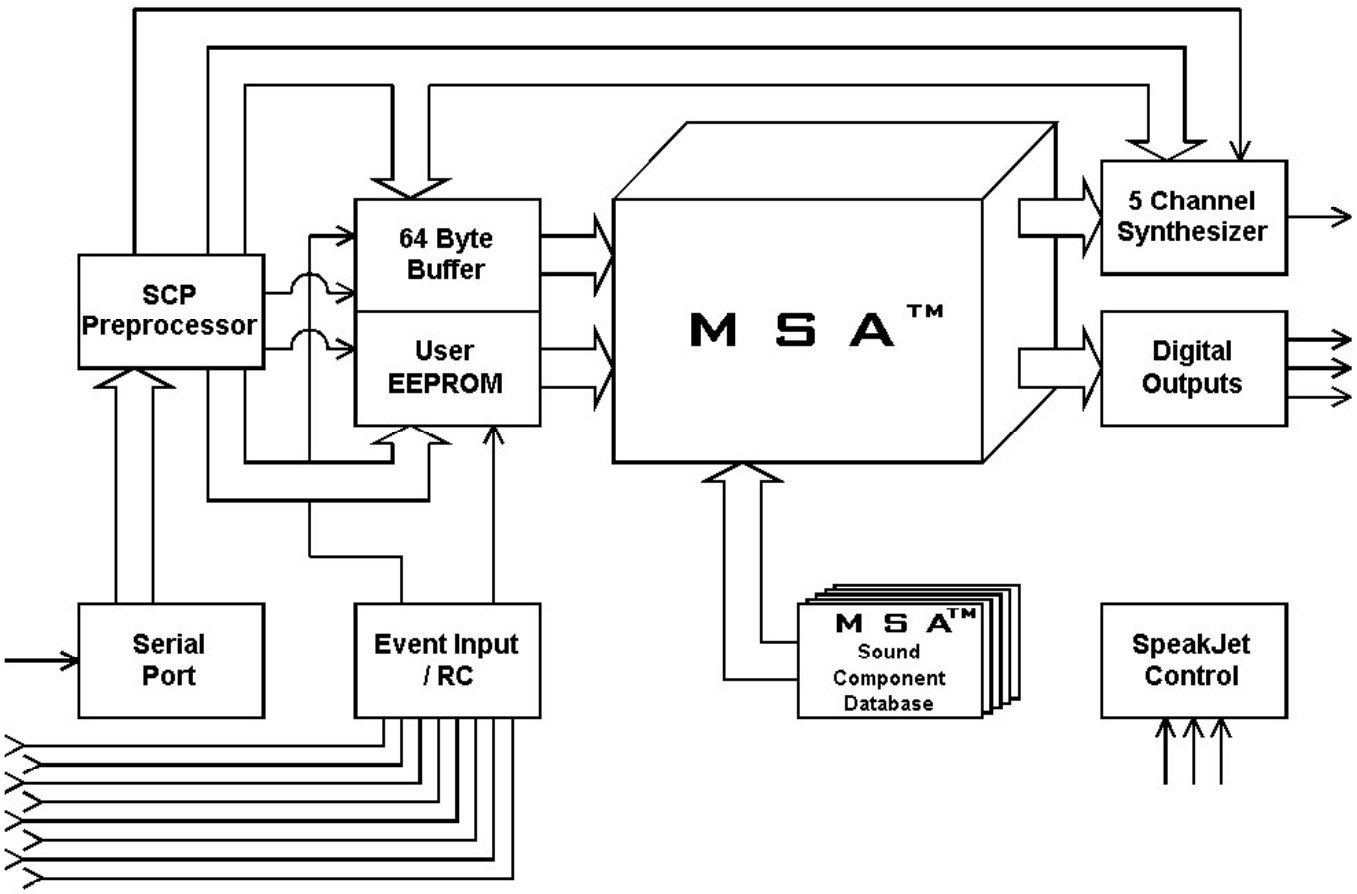


Figure 2. SpeakJet Block Diagram

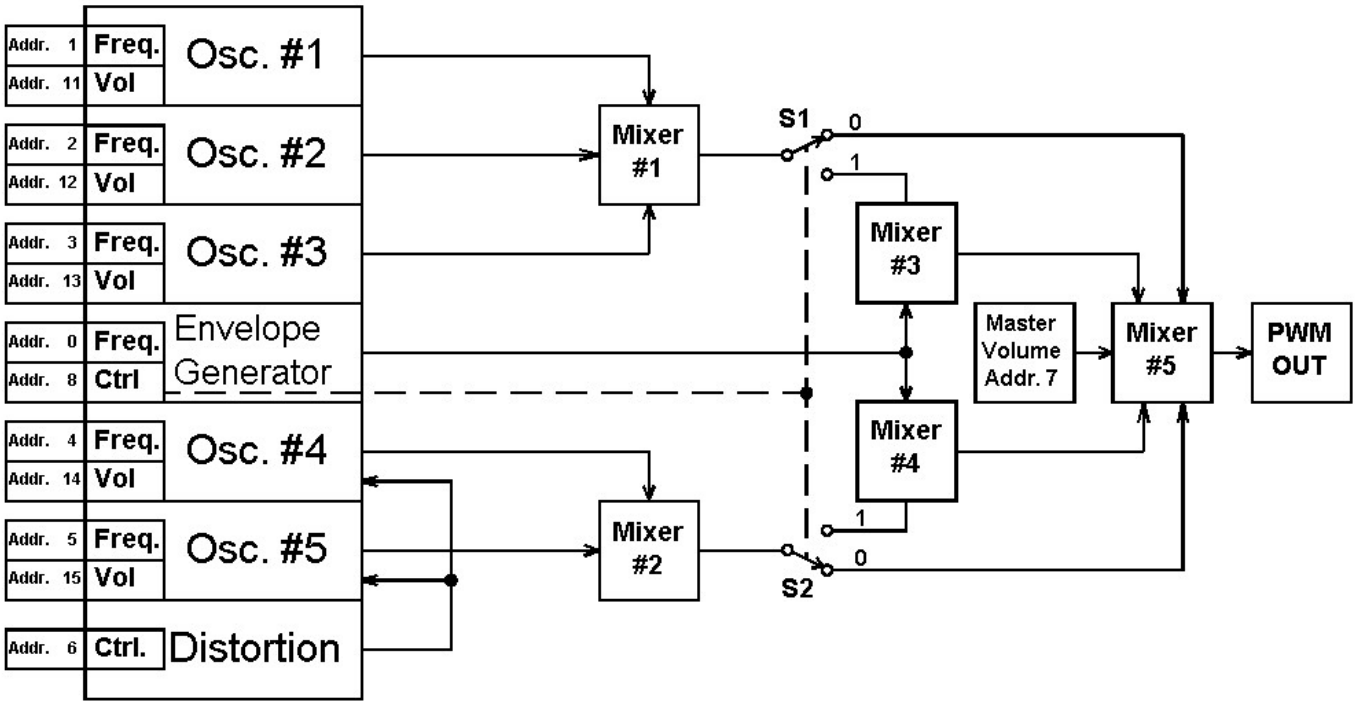


Figure 3. SpeakJet 5 Channel Synthesizer Block Diagram

Preliminary Data

Preliminary Data

Phoneme/Sound Effect Code List:

Code	Phoneme	Samle Words	Msec.	Phoneme Type
0	P0		0	Silence
1	P1		100	Silence
2	P2		200	Silence
3	P3		700	Silence
4	P4		10	Silence /w Trans
5	P5		30	Silence /w Trans
6	P6		60	Silence /w Trans
128	IY	See, Even, Feed	70	Voiced Long Vowel
129	IH	Sit, Fix, Pin	70	Voiced Short Vowel
130	EY	Hair, Gate, Beige	70	Voiced Long Vowel
131	EH	Met, Check, Red	70	Voiced Short Vowel
132	AY	Hat, Fast, Fan	70	Voiced Short Vowel
133	AX	Cotten	70	Voiced Short Vowel
134	UX	Luck, Up, Uncle	70	Voiced Short Vowel
135	OH	Hot, Clock, Fox	70	Voiced Short Vowel
136	AW	Father, Fall	70	Voiced Short Vowel
137	OW	Comb, Over, Hold	70	Voiced Long Vowel
138	UH	Book, Could, Should	70	Voiced Short Vowel
139	UW	Food, June	70	Voiced Long Vowel
140	MM	Milk, Famous,	70	Voiced Nasal
141	NE	Nip, Danger, Thin	70	Voiced Nasal
142	NO	No, Snow, On	70	Voiced Nasal
143	NGE	Think, Ping	70	Voiced Nasal
144	NGO	Hung, Song	70	Voiced Nasal
145	LE	Lake, Alarm, Lapel	70	Voiced Resonate
146	LO	Clock, Plus, Hello	70	Voiced Resonate
147	WW	Wool, Sweat	70	Voiced Resonate
148	RR	Ray, Brain, Over	70	Voiced Resonate
149	IYRR	Clear, Hear, Year	200	Voiced R Color Vowel
150	EYRR	Hair, Stair, Repair	200	Voiced R Color Vowel
151	AXRR	Fir, Bird, Burn	190	Voiced R Color Vowel
152	AWRR	Part, Farm, Yarn	200	Voiced R Color Vowel
153	OWRR	Corn, Four, Your	185	Voiced R Color Vowel
154	EYIY	Gate, Ate, Ray	165	Voiced Diphthong
155	OHIY	Mice, Fight, White	200	Voiced Diphthong
156	OWIY	Boy, Toy, Voice	225	Voiced Diphthong
157	OHIH	Sky, Five, I	185	Voiced Diphthong
158	IYEH	Yes, Yarn, Million	170	Voiced Diphthong
159	EHLL	Saddle, Angle, Spell	140	Voiced Diphthong
160	IYUW	Cute, Few,	180	Voiced Diphthong
161	AXUW	Brown, Clown, Thousand	170	Voiced Diphthong
162	IHWW	Two, New, Zoo	170	Voiced Diphthong
163	AYWW	Our, Ouch, Owl	200	Voiced Diphthong
164	OWWW	Go, Hello, Snow	131	Voiced Diphthong
165	JH	Dodge, Jet, Savage	70	Voiced Affricate
166	VV	Vest, Even,	70	Voiced Fictive
167	ZZ	Zoo, Zap	70	Voiced Fictive
168	ZH	Azure, Treasure	70	Voiced Fictive
169	DH	There, That, This	70	Voiced Fictive
170	BE	Bear, Bird, Beed	45	Voiced Stop
171	BO	Bone, Book Brown	45	Voiced Stop
172	EB	Cab, Crib, Web	10	Voiced Stop
173	OB	Bob, Sub, Tub	10	Voiced Stop
174	DE	Deep, Date, Divide	45	Voiced Stop
175	DO	Do, Dust, Dog	45	Voiced Stop
176	ED	Could, Bird	10	Voiced Stop
177	OD	Bud, Food	10	Voiced Stop
178	GE	Get, Gate, Guest,	55	Voiced Stop
179	GO	Got, Glue, Goo	55	Voiced Stop
180	EG	Peg, Wig	55	Voiced Stop
181	OG	Dog, Peg	55	Voiced Stop
182	CH	Church, Feature, March	70	Voiceless Affricate
183	HE	Help, Hand, Hair	70	Voiceless Fricative
184	HO	Hoe, Hot, Hug	70	Voiceless Fricative
185	WH	Who, Whale, White	70	Voiceless Fricative
186	FF	Food, Effort, Off	70	Voiceless Fricative
187	SE	See, Vest, Plus	40	Voiceless Fricative

Table A

Phoneme/Sound Effect Code List Continued:

Code	Phoneme	Sample Words	MS	Phoneme Type
188	SO	So, Sweat	40	Voiceless Fricative
189	SH	Ship, Fiction, Leash	50	Voiceless Fricative
190	TH	Thin, month	40	Voiceless Fricative
191	TT	Part, Little, Sit	50	Voiceless Stop
192	TU	To, Talk, Ten	70	Voiceless Stop
193	TS	Parts, Costs, Robots	170	Voiceless Stop
194	KE	Can't, Clown, Key	55	Voiceless Stop
195	KO	Comb, Quick, Fox	55	Voiceless Stop
196	EK	Speak, Task	55	Voiceless Stop
197	OK	Book, Took, October	45	Voiceless Stop
198	PE	People, Computer	99	Voiceless Stop
199	PO	Pow, Copy	99	Voiceless Stop
200	R0		80	Robot
201	R1		80	Robot
202	R2		80	Robot
203	R3		80	Robot
204	R4		80	Robot
205	R5		80	Robot
206	R6		80	Robot
207	R7		80	Robot
208	R8		80	Robot
209	R9		80	Robot
210	A0		300	Alarm
211	A1		101	Alarm
212	A2		102	Alarm
213	A3		540	Alarm
214	A4		530	Alarm
215	A5		500	Alarm
216	A6		135	Alarm
217	A7		600	Alarm
218	A8		300	Alarm
219	A9		250	Alarm
220	B0		200	Beeps
221	B1		270	Beeps
222	B2		280	Beeps
223	B3		260	Beeps
224	B4		300	Beeps
225	B5		100	Beeps
226	B6		104	Beeps
227	B7		100	Beeps
228	B8		270	Beeps
229	B9		262	Beeps
230	C0		160	Biological
231	C1		300	Biological
232	C2		182	Biological
233	C3		120	Biological
234	C4		175	Biological
235	C5		350	Biological
236	C6		160	Biological
237	C7		260	Biological
238	C8		95	Biological
239	C9		75	Biological
240	D0	0	95	DTMF
241	D1	1	95	DTMF
242	D2	2	95	DTMF
243	D3	3	95	DTMF
244	D4	4	95	DTMF
245	D5	5	95	DTMF
246	D6	6	95	DTMF
247	D7	7	95	DTMF
248	D8	8	95	DTMF
249	D9	9	95	DTMF
250	D10	*	95	DTMF
251	D11	#	95	DTMF
252	M0	Sonar Ping	125	Miscellaneous
253	M1	Pistol Shot	250	Miscellaneous
254	M2	WOW	530	Miscellaneous
255	NS		0	Silence

Table A

Preliminary Data

Control Codes:

Dec.	Character	ASCII Use	SpeakJet Use
000	NUL	(Null char.)	Pause 0
001	SOH	(Start of Header)	Pause 1
002	STX	(Start of Text)	Pause 2
003	ETX	(End of Text)	Pause 3
004	EOT	(End of Transmission)	Pause 4
005	ENQ	(Enquiry)	Pause 5
006	ACK	(Acknowledgment)	Pause 6
007	BEL	(Bell)	Play Next Sound Fast
008	BS	(Backspace)	Play Next Sound Slow
009	HT	(Horizontal Tab)	-
010	LF	(Line Feed)	-
011	VT	(Vertical Tab)	-
012	FF	(Form Feed)	-
013	CR	(Carriage Return)	-
014	SO	(Shift Out)	Play Next Sound High
015	SI	(Shift In)	Play Next Sound Low
016	DLE	(Data Link Escape)	Wait
017	DC1	(XON)(Device Control 1)	-
018	DC2	(Device Control 2)	Play Next Sound Soft
019	DC3	(XOFF)(Device Control 3)	-
020	DC4	(Device Control 4	Volume, X
021	NAK	(Negative Acknowledgement)	Speed, X
022	SYN	(Synchronous Idle)	Pitch, X
023	ETB	(End of Trans. Block)	Bend, X
024	CAN	(Cancel)	PortCtr, X
025	EM	(End of Medium)	Port, X
026	SUB	(Substitute)	Repeat, X
027	ESC	(Escape)	-
028	FS	(File Separator)	Call Phrase, X
029	GS	(Group Separator)	Goto Phrase, X
030	RS	(Request to Send)	Delay, X
031	US	(Unit Separator)	Reset Defaults
032 - to - 127			Reserved For TTS
128 - to - 253			126 Sound codes
254			Silence
255			No Sound / End of Phrase.

Control Codes Details:

0 - 6 = Pauses.
Pauses of various durations. These will cause the volume to ramp down, wait a specified amount of time and the ramp back up. 1, 2 & 3, ramp the volume while the format frequencies are being changed. 4, 5 & 6 wait for silence before changing the format frequencies.

0 = 0ms	1 = 100ms	2 = 200ms	3 = 700ms
4 = 30ms	5 = 60ms	6 = 90ms	

7 = Fast.
Plays the next phoneme at 1/2 the time it normally would play.

8 = Slow
Plays the next phoneme at 1 and 1/2 the time it normally would play.

14 = Stress.
Plays the next phoneme with a small amount of stress in the voice.

15 = Relax
Plays the next phoneme with a small amount of relaxation in the voice.

16 = Wait
This command will stop the voicing and wait for a start command. The Start command can be issued by either sending the SCP start command or by changing the state of one of the input lines that has been previously set to do a Start.

18 = Soft
Plays the next sound with a reduced volume of 50%

20 = Volume, X
This command sets the master volume level. A value will need to be sent after the volume command that specifies the desired volume. Volume levels can range from 0 to 127. The default is 96.

21 = Speed, X
This command sets the play speed. A value will need to be sent after the speed command that specifies the desired speed. Speeds can range from 0 to 127. The default is 114.

22 = Pitch, X
This command sets the Vocalization Pitch in Hertz. A value will need to be sent after the pitch command that specifies the desired pitch. The vocalization pitch is what makes a voice sound High pitched or Low pitched. For singing, the pitch has a range of 3 full octaves (32Hz to 240hz). The Vocalization Pitch works only on sounds that are voiced. Pitches can range from 0 to 255. The default is 88. Note that anything under 30 starts to sound like clicks instead of a voice. Also Note that a value of 0 = 0 Hz and thusly, will not actually vocalize.

23 = Bend, X
This command sets the frequency Bend. A value will need to be sent after the Bend command that specifies the desired Bend. The frequency Bend adjusts the output frequencies of the oscillators. This will change the voicing from a deep-hollow sounding voice to a High-metallic sounding voice. Bends can range from 0 to 15. The default is 5.

24 = PortCtr, X
This command sets the Port Control Value. A value will need to be sent after the PortCtr command that specifies the desired function of the output lines. The Output line control bits are binaurally encoded where a 1 indicates that the output function is chip controlled and a 0 indicates that the output function is user controlled. Bit 0 corresponds to OUT0, etc… PortCtr values can range from 0 to 7. The default is 7.

25 = Port, X
This command sets the Port Output Value. A value will need to be sent after the Port command that specifies the desired state of the output lines. When the Output line control bits are set to 0, the corresponding port bit is represented on the output line. Bit 0 corresponds to OUT0, etc… Port values can range from 0 to 7. The default is 0.

26 = Repeat, X
This command sets a number of times to Repeat the next code. A value will need to be sent after the Reset command that specifies the number of times to repeat the next command.
The Repeat range is from 0 to 255.

28 = Call Phrase, X
This command specifies which EEPROM phrase to play then to return from.
This can be nested 3 deep maximum.

29 = Goto Phrase, X
This command specifies which EEPROM phrase to play.

30 = Delay, X
This command specifies the number of 10ms intervals to delay before continuing on to the next code. The Delay range is from 0 to 255.

31 = Reset
This command resets the Volume, Speed, Pitch and Bend to the default values.

Preliminary Data